A statistical test of emission from unresolved point sources

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **405,** 1777–1786 (2010) doi:10.1111/j.1365-2966.2010.16550.x

A statistical test of emission from unresolved point sources

Tracy R. Slatyer[1,2]★ and Douglas P. Finkbeiner[1,2]

[1]*Physics Department, Harvard University, Cambridge, MA 02138, USA*
[2]*Institute for Theory and Computation, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*

Accepted 2010 February 18. Received 2010 January 28; in original form 2009 October 22

ABSTRACT

We describe a simple test of the spatial uniformity of an ensemble of discrete events. Given an estimate for the point-source luminosity function and an instrumental point spread function, a robust upper bound on the fractional point-source contribution to a diffuse signal can be found. We verify with Monte Carlo tests that the statistic has advantages over the two-point correlation function for this purpose, and derive analytic estimates of the statistic's mean and variance as a function of the point-source contribution. As a case study, we apply this statistic to recent gamma-ray data from the *Fermi* Large Area Telescope, and demonstrate that at energies above 10 GeV the contribution of unresolved point sources to the diffuse emission is small in the region relevant for study of the *WMAP* Haze.

Key words: methods: statistical – gamma-rays: diffuse background.

1 INTRODUCTION

Statistical tests of isotropy have a long history in astronomy. A common question is 'what fraction of the observed emission could originate from unresolved point sources?' For example, possible point-source contributions to the extragalactic X-ray background were investigated by Scheuer (1974), and the small-angle power spectrum of the cosmic far-infrared background has been used to estimate the isotropic component (Kashlinsky, Mather & Odenwald 1996). More recently, the Auger team tested the isotropy of ultrahigh energy cosmic ray events by cross-correlating with positions of known active galactic nuclei (AGN; Abraham et al. 2007, 2008) to provide information on their origin. However, in the absence of an appropriate external catalog such cross-correlation methods cannot be used, motivating consideration of a more general approach.

In some cases, a detector provides binned counts (e.g. pixels in a CCD); in other cases, photon event directions are reconstructed in some other way (e.g. a gamma-ray pair conversion telescope). In the latter case, it is desirable to apply statistics that do not require binning of the data, as binning introduces additional arbitrary parameters into the problem. In the limit of low flux density, where the mean density of photon events (hereafter 'events') is much less than one per point spread function (PSF), explicit detection of point sources may become impractical and estimation of the unresolved point-source flux becomes especially difficult. In some cases, the two-point correlation function, or some modified form (e.g. Ave et al. 2009), is used as a test. However, the Fourier transform of a field of point sources has significant phase correlation, and a two-point function (or a power spectrum) discards this phase information. Higher order correlation statistics capture it, but are somewhat cumbersome to use. In the following, we describe a statistic that is easy to understand and evaluate and is optimized to address this question, particularly in the case of fairly sparse data sets with (on average) $\lesssim$1 event per PSF circle.

The key insight is that if a substantial fraction of the photons come from point sources, it is much more likely that two photons appear within one PSF of each other than in the diffuse case. This is true even if the expected integrated flux is of the order of one count. Likewise, the number of PSF circles containing *no* counts is larger if point sources contribute. These considerations motivate us to define a ratio between the fraction of 'isolated events' and the fraction of 'empty circles.' This ratio is very closely related to the fraction of diffuse emission and can be calibrated with Monte Carlo simulations for specific choices of instrumental parameters and a putative luminosity function. The two-point function, in contrast, is weighted by the density squared and is *not* proportional to the desired quantity.

In the following sections, we define the statistic, estimate its variance, show how it behaves in various limits, generalize it to the case where many events appear in every PSF circle and show a practical application to recently released data from the *Fermi* Gamma-ray Space Telescope.

2 DEFINITION OF THE STATISTIC

For each event, we consider the number of neighbouring events within some test radius r; the natural choice of r is determined by the PSF of the detector. Let the fraction of events with zero neighbours (isolated events) be denoted as n_I. Now, consider a random distribution of points within the signal region and the number of events within r of each of these points. Let the fraction of points with zero events within radius r (empty circles) be denoted as n_E

★E-mail: tslatyer@fas.harvard.edu

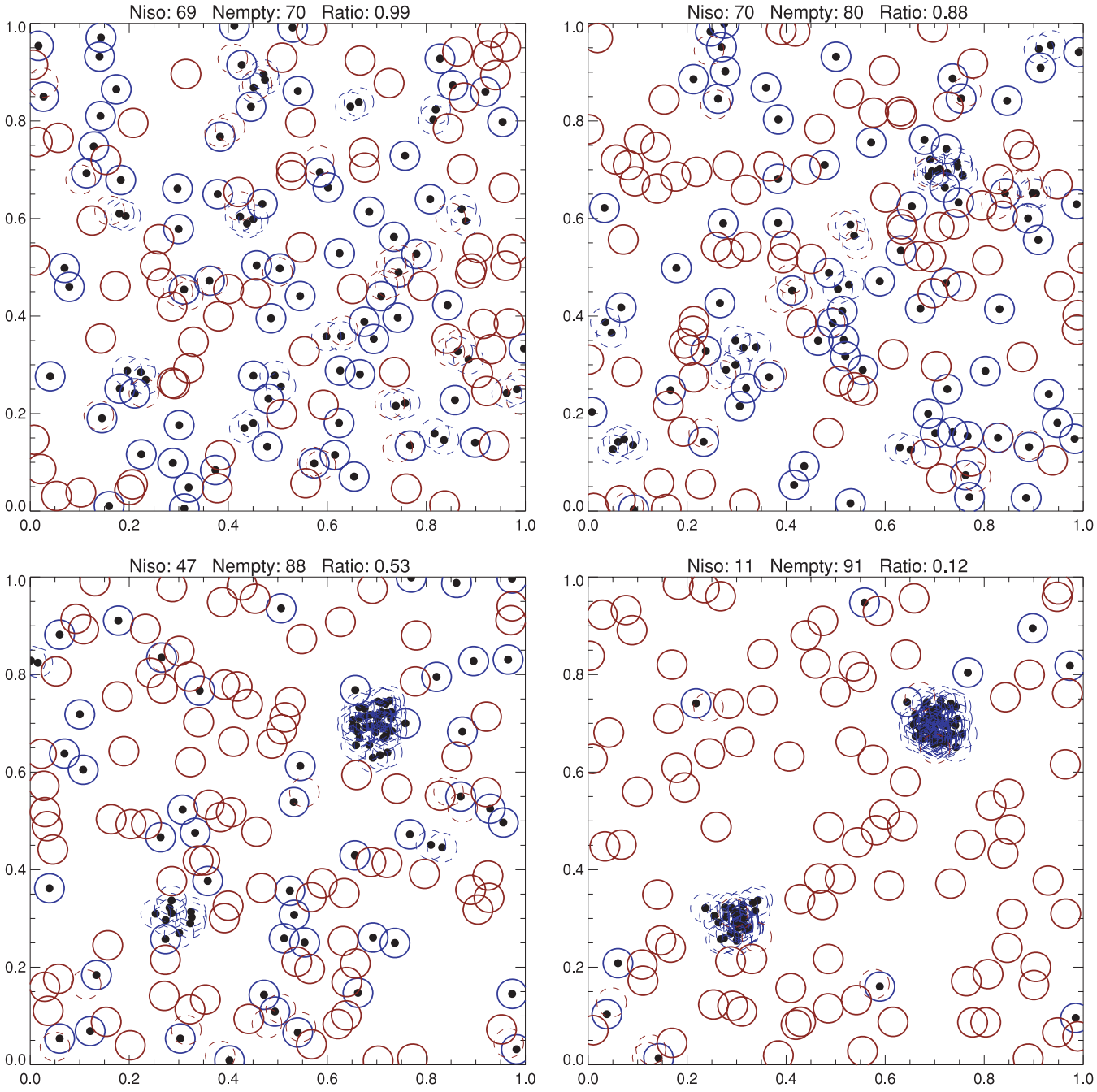

Figure 1. In each panel, photon events (dots) are either isolated (solid blue circles) or not (dashed blue circles). Random circles are either empty (solid red) or not(dashed red). In each case, 100 events and 100 random circles are shown, so the ratio, R, can be visualized here as the number of solid blue circles divided by the number of solid red circles. In practise, one uses a large number of random circles to reduce noise. The panels contain either no point sources (upper-left panel), or 15 per cent (upper-right panel), 50 per cent (lower-left panel), or 90 per cent (lower-right panel) point-source flux in sources located at (0.3,0.3) and (0.7,0.7).

(the number of randomly distributed points should be as large as is computationally feasible, to reduce the Poisson error in the fraction n_E).

Fig. 1 illustrates this idea. If the signal photons are randomly distributed, as in the case of a uniform diffuse signal, $n_I \approx n_E$ (up to Poisson fluctuations). If the events are clustered on the scale r, on the other hand, n_I falls (as more events have close neighbours) and n_E rises (as the events are clumped, more of the signal region contains no events at all).

If a large fraction of the photon counts originate from unresolved point sources, we expect significant correlations between photon positions, even if the expected counts from a point source are $\lesssim 1$. Consequently, the ratio $R = n_I/n_E$ serves as a simple measure of the fraction of flux originating from point sources as opposed to uniform diffuse emission.

Larger-scale non-uniformities, such as a gradient in the distribution of diffuse photons, have only a small effect on R, as we show explicitly with Monte Carlo tests in Section 6. In regions of

high event density both n_I and n_E are suppressed, so the suppression largely cancels out in the ratio. However, such a gradient *will* tend to slightly lower R in the diffuse limit, since the points relevant to computing n_E are uniformly distributed, whereas those relevant to computing n_I are concentrated in regions of high flux (this in turn can reduce the gradient of R with respect to the fraction of diffuse flux, since in the point-source-dominated limit R is small but independent of any inhomogeneity in the diffuse emission). The point-source emission derived from R is thus best interpreted as an upper bound, in cases where the diffuse emission is suspected to have significant spatial variation.

3 ANALYTIC ESTIMATE FOR THE MEAN AND VARIANCE

To understand the behaviour of the ratio R as a function of the diffuse emission fraction, consider a related but simpler problem, where we treat the signal region as a grid and count the number of events in each cell. In this case, an 'isolated' event is one with no other events in the same cell and R is simply the ratio of the fraction of events which are isolated to the fraction of cells which are empty. The total number of cells is N, n of which are empty and m of which have a single event.

The probability of any given cell being empty is p_0 and containing a single event is p_1. That is, $\langle n\rangle = p_0 N$ and $\langle m\rangle = p_1 N$. The joint probability of having precisely n empty cells and m single-event cells is

$$P(m,n) = p_0^n p_1^m (1-p_0-p_1)^{N-n-m}\frac{N!}{n!m!(N-m-n)!}. \quad (1)$$

The fraction of isolated events is $\langle n_I\rangle = m/N_{\rm event}$, while the empty fraction is $\langle n_E\rangle = n/N$. In this case, $R = m/(n+1)\times N/N_{\rm event}$. As we will see, $\langle m/(n+1)\rangle$ is an unbiased estimator of the probability ratio p_1/p_0, which corresponds to R as defined above in the limit of large n, and is equivalent to the fraction of the flux originating from diffuse emission if the point sources are sufficiently bright.

The expectation value of $m/(n+1)$ is given by

$$\left\langle \frac{m}{n+1}\right\rangle = \sum_{n=0}^{N-1}\sum_{m=1}^{N-n}\frac{m}{n+1}P(m,n) = \frac{p_1}{p_0}\left[1-(1-p_0)^N\right]. \quad (2)$$

The expectation value $\langle (m/(n+1))^2\rangle$ is given by

$$\begin{aligned}&\sum_{n=0}^{N-1}\sum_{m=1}^{N-n}\left(\frac{m}{n+1}\right)^2 P(m,n)\\ &= \frac{p_1}{(-1+p_0)^2 p_0}\Bigg\{-\left[-1+(1-p_0)^N\right](-1+p_0)p_1\\ &\quad + N(1-p_0)^N p_0(1-p_0+Np_1)\\ &\quad \times {}_3F_2\left(\{1,1,1-N\},\{2,2\},\frac{p_0}{-1+p_0}\right)\Bigg\}\\ &\to \left(\frac{p_1}{p_0}\right)^2 + \frac{p_1}{Np_0^2}\left(1+\frac{p_1}{p_0}\right), \quad N\to\infty. \end{aligned} \quad (3)$$

Consequently, neglecting terms suppressed by a large power of $(1-p_0)$, the mean μ and variance σ^2 of $m/(n+1)$ in the case of large N are given by

$$\mu = p_1/p_0, \quad \sigma^2 = p_1(1+p_1/p_0)/Np_0^2. \quad (4)$$

The corresponding quantities for R are obtained by a rescaling by $N/N_{\rm event}$ and $(N/N_{\rm event})^2$, respectively:

$$\langle R\rangle = \frac{p_1}{p_0}\frac{N}{N_{\rm event}}, \quad \left(\frac{\sigma_R}{\langle R\rangle}\right)^2 = \frac{1+p_1/p_0}{Np_1}\sim\frac{1}{m}. \quad (5)$$

It is unsurprising that the fractional uncertainty in R is approximately $1/\sqrt{m}$, where m is the number of isolated events.

In the limit where all the flux is diffuse with mean rate λ, then $p_1 = \lambda e^{-\lambda}$, $p_0 = e^{-\lambda}$, and $\langle N_{\rm event}\rangle = N\lambda$, so R has mean 1 and variance $(1+\lambda)e^{\lambda}/N\lambda$. Now consider the addition of point sources. We can approximate the effect of adding point sources by choosing T cells which each gain >1 events. Let the total number of added events be Γ. Then on average a fraction T/N of the isolated events will no longer be isolated, and a fraction T/N of the empty cells will no longer be empty (by assuming each cell that gains events gains >1, we ensure that empty cells will not gain isolated events).

Thus, when point sources are added both p_0 and p_1 are multiplied by $(1-T/N)$ and $N_{\rm event}$ becomes $N\lambda+\Gamma$. Thus, the mean of R becomes $N\lambda/(N\lambda+\Gamma)$ = diffuse flux / total flux and the variance becomes

$$\begin{aligned}\sigma^2(R) &= \frac{1}{N\lambda+\Gamma}\left[\frac{(1+\lambda)\,e^{\lambda}}{(1+\Gamma/N\lambda)(1-T/N)}\right]\\ &= \frac{(1+\lambda)\,e^{\lambda}}{\text{total counts}}\\ &\quad\times\left(\frac{\text{diffuse flux/total flux}}{\text{fraction of pixels with no point sources}}\right).\end{aligned} \quad (6)$$

So, in this simple model the mean of the ratio R is precisely the diffuse flux divided by the total flux. The variance in R is generally Poisson, scaling as 1/(total photon counts) provided the number of counts per pixel from diffuse emission is $\lesssim 1$, but is reduced when the diffuse flux is much smaller than the total flux. (It is possible for the diffuse flux to be much less than the total flux even when most cells do not contain point sources, but not the converse, unless λ is rather large, in which case the sample size or the cell size should be reduced.) The variance grows rapidly as λ becomes greater than 1: if the cell size (corresponding to angular resolution) is too large relative to the amount of diffuse emission, R is not a good measure of the fraction of diffuse flux (however, see Section 3.1).

If there is a significant contribution from weak point sources that may only add a single isolated event to an empty cell (or alternatively, if the cells are sufficiently small that this is common even for stronger point sources), then the mean of R is no longer given simply by the fraction of diffuse flux. For example, suppose S cells gain exactly one event when point sources are added, and T cells in total are affected by the addition of point sources (so $T\geq S$): then, the addition of point sources sends $p_0\to(1-T/N)p_0$ and $p_1\to(1-T/N)p_1+(S/N)p_0$. Then, we obtain

$$\langle R\rangle = \frac{N}{N\lambda+\Gamma}\left(\lambda+\frac{S}{N-T}\right), \quad (7)$$

$$\sigma^2(R) = \left(\frac{1}{N\lambda+\Gamma}\right)\frac{e^{\lambda}\left(1+\lambda+\frac{S}{N-T}\right)\left(1+\frac{S}{\lambda(N-T)}\right)}{(1+\Gamma/N\lambda)\,(1-T/N)}. \quad (8)$$

If the number of cells containing single one-photon point sources is small compared to the number of cells unaffected by point sources ($S\ll N-T$), and also the number of photons from these weak point sources is small compared to the number of diffuse events from cells unaffected by point sources [$S\ll\lambda\,(N-T)$], we recover the previous result. If the second condition fails to hold (as occurs in the limit of low diffuse emission, independent of the point-source luminosity function) then $\langle R\rangle$ asymptotes to (flux from isolated one-photon point sources)/(total flux) $\times$ $1/(1-T/N)$, as $\lambda\to 0$ with the total flux held constant. If this limiting value is $\gtrsim 1$ then R has no discriminatory power, and the situation is not improved by higher statistics: this is simply the statement that there is no difference

between diffuse emission and a very large number of very faint uniformly distributed point sources.

3.1 Extension to the case of large λ

One region of parameter space in which this test breaks down is where $\lambda \geq 1$. However, a simple generalization of the statistic can be useful in this case. Suppose we make a histogram of the number of nearest neighbours each event possesses and measure the peak of the histogram to be some number of neighbours $n_{\rm crit}$. Let us redefine n_I as the fraction of events with $n_{\rm crit}$ or fewer neighbours and n_E as the fraction of points with $n_{\rm crit}$ or fewer events within the test radius r. Then, for Monte Carlo realizations of diffuse flux plus some randomly distributed point sources we can again measure the ratio n_I/n_E and use it as a measure of how correlated the photon events are (see Section 6.2).

4 THE POINT-SOURCE LUMINOSITY FUNCTION

Clearly, the sensitivity of the test depends critically on the fraction of point sources which contribute at most one count to the data, which is determined by the point-source luminosity function. In the limit where the point-source flux is dominated by (a very large number of) uniformly distributed sources which each produce an average number of counts $\ll 1$, unresolved point sources are practically indistinguishable from diffuse emission.

In the Monte Carlo tests which follow, we treat the luminosity function as some unbroken power law $\mathrm{d}N/\mathrm{d}S = S^{-\alpha}$ between integrated flux limits $S_{\rm min}$ and $S_{\rm max}$. $S_{\rm max}$ is bounded above by the faintest point sources which can be resolved and masked out. $S_{\rm min}$, on the other hand, is not determined by known properties of the experiment, and limits derived with a particular value of $S_{\rm min}$ should be interpreted as placing a limit on the contribution from point sources with average luminosity above $S_{\rm min}$.

The effect of $S_{\rm min}$ on the behaviour of R depends on the value of the spectral index α: if $\alpha > 2$ then most of the flux originates from the faintest point sources, and the total flux diverges at low luminosity. In this case, the power law generally breaks to a much shallower slope at some low luminosity, meaning that most of the flux originates from point sources with luminosities close to the break: we can approximate this behaviour simply by cutting off the luminosity function at $S_{\rm min}$. If $\alpha < 2$ then most of the flux is concentrated in the brighter point sources and changing $S_{\rm min}$ has little effect on R.

To estimate the relevant range of α, we examine studies of known populations of point sources. The gamma-ray luminosity functions of AGN contributing to the extragalactic gamma-ray background have been studied using the first three months of data from the *Fermi* Large Area Telescope (LAT; Abdo et al. 2009a). The luminosity function for BL Lac objects was found to be well described by a single power law with $\alpha = 2.17 \pm 0.05$. The luminosity function for flat spectrum radio quasars (FSRQs) was well described by a power law with $\alpha = 2.58 \pm 0.19$ at high redshifts ($z \geq 1$), indicating that at high redshifts the *Fermi* LAT is sampling the bright (steep) end of the luminosity function, but at $z \leq 1$ the best-fitting value of the slope was $\alpha = 1.56 \pm 0.10$. The X-ray luminosity functions of the same classes of objects have been studied by Padovani et al. (2007), with BL Lac objects measured to have $\alpha = 2.12 \pm 0.16$ and $\alpha = 1.6 - 1.9$ for FSRQs at $z \leq 1$.

The X-ray luminosity function of high-mass X-ray binaries (HMXBs) was found by Grimm, Gilfanov & Sunyaev (2003) to have slope $\alpha = 1.61 \pm 0.12$. Studies of low-mass X-ray binaries (LMXBs) in Centaurus A (Voss & Gilfanov 2006; Voss et al. 2009) yield a slope of $\alpha = 1.8 - 2.0$ at high luminosity, flattening to $\alpha \sim 1.2$ at low luminosity, in agreement with earlier studies of LMXBs (Gilfanov 2004; Kim & Fabbiano 2004).

Table 1. Benchmark parameters for luminosity functions $\mathrm{d}N/\mathrm{d}S \sim S^{-\alpha}$ with $S_{\rm min} < S < S_{\rm max}$.

	α	$S_{\rm max}$	$S_{\rm min}$
Benchmark 1	2.2	10	0.1
Benchmark 2	1.8	100	1

In the examples in the following section, therefore, we take our two benchmark models to have $\alpha = 1.8, 2.2$, and also demonstrate the effect of varying α between 1.5 and 3.0. Smaller values for α improve the ability of our statistic to distinguish point sources from diffuse emission, simply because more of the flux originates from bright sources.

5 MONTE CARLO EXAMPLES

To examine the usefulness of the ratio R as a measure of the point-source flux, we employ a Monte Carlo approach. We consider a given angular 'signal region' and a smaller 'signal window' within that region (the purpose of this distinction is to eliminate edge effects). Within the signal region, we generate a uniform random distribution of point-source locations, for a certain number of sources (which we vary, as a proxy for varying the total flux from point sources). The expected number of counts for each point source is drawn from a power-law distribution with spectral index α, with cutoffs at a minimum expected number of counts $S_{\rm min}$ and a maximum expected number of counts $S_{\rm max}$. For our benchmark models, we take (1) a 'pessimistic' set of parameters $\alpha = 2.2$, $S_{\rm max} = 10$, $S_{\rm min} = 0.1$ and (2) an 'optimistic' set of parameters $\alpha = 1.8$, $S_{\rm max} = 100$, $S_{\rm min} = 1$ (see Table 1). Once the expected counts from each source have been obtained, the number of counts actually observed from each source is determined by a Poisson draw; their angular distribution is determined by the detector PSF, which for simplicity we model as a Gaussian with standard deviation σ.

We compute the total number of counts lying within the signal window and originating from the point sources, and subtract this quantity from the desired total number of counts in the signal window. The result approximates the number of diffuse events we wish to generate in the signal window. We then generate a uniform random distribution of diffuse events in the signal region, with the total number of events given by the desired diffuse counts in the signal window rescaled to the greater area of the signal region. The total photon distribution in the signal window (diffuse + point sources) then has approximately the correct number of events. We can also incorporate non-uniformities in the diffuse flux at this point.

Having produced our test data, we need to compute the ratio R. For each event within the signal window, we use the publicly available IDL routine SPHEREMATCH[1] to find its neighbours in the set of events in the larger signal region. This eliminates edge effects, i.e. a spuriously high number of isolated events at the edges of the signal window. We repeat the process for a random distribution of points within the signal window, to compute n_E, and plot the resulting R

[1]The IDL routines used in this analysis are available as part of v5_4_8 of the IDLUTILS product at http://sdss3data.lbl.gov/software/idlutils.

against the fraction of flux in the signal window originating from point sources.

For these examples, we take the default signal window to be the region $|l| < 15$, $|b| < 15$, the PSF to be $0.^{\circ}2$, and the total number of counts to be 3000, corresponding to a mean $\lambda = 0.4$ events per PSF circle. Note that the statistic is insensitive to the shape of the signal window, and if we rescale the PSF by some factor a and the area of the signal window by a^2, and hold the number of events constant, then this is just equivalent to a unit redefinition and does not change the results.

When we refer to the 'PSF', we mean the standard error 1σ. Optical astronomers often use the full-width at half maximum (FWHM), which for a Gaussian equals 2.355σ and contains 50 per cent of the flux. Gamma-ray astronomers often use the radius of 68 or 95 per cent containment (1.51σ or 2.45σ, respectively). The R statistic is not greatly sensitive to the exact choice of r. We will demonstrate the effect of varying the PSF while holding the other parameters constant (thus changing λ, the mean number of events per PSF circle), and of increasing the number of counts while holding λ constant, by increasing the signal window. We will also generally assume that the test radius r is equal to 1σ, but show the effect of using a different test radius, for both benchmark luminosity functions.

In the plots that follow, we will describe the 'sensitivity' of this test by two representative measures.

(i) The maximum value of the point-source flux fraction consistent with $R = 1$, within the 95 per cent confidence limits.

(ii) Bounds on the point-source flux fraction when the true point-source contribution is half the signal, obtained by averaging the limits which would be obtained from an R-measurement over the histogram of R-values produced in this scenario.

The statistic is most powerful where the bound in (i) is small and the limits in (ii) are close together. If either the variance of R (for fixed point-source flux) becomes large or the mean of R is slowly varying with respect to the point-source flux fraction, then both of these measures will blow up. We also overplot the sensitivity estimates obtained from the analytic approximation for R in the grid model described in Section 3, using the relations in Appendix A, for comparison to the results of the MC runs.

6 RESULTS

We find a strong linear relationship between R and the fraction of the flux due to diffuse emission, as shown in Fig. 2 for the two benchmark sets of parameters. As expected, the 'optimistic' benchmark parameters render R more sensitive to the fraction of diffuse emission.

6.1 Dependence on the point-source luminosity function and the effect of improved statistics

Fig. 3 shows the effect of varying the spectral index of the point-source luminosity function, for the two benchmark choices of $S_{\rm min}$, $S_{\rm max}$. As expected, smaller values of α yield better performance for R as a precise estimator of the diffuse flux. For the pessimistic benchmark model, when $\alpha \sim 3.0$ the method has lost most of its discriminatory power: however, for more well-motivated power laws $\alpha \lesssim 2.5$, useful limits can still be obtained. For the optimistic benchmark model, R can be used to place strong limits on the diffuse emission even for very steep power laws $\alpha \sim 3.0$, largely because the assumed break in the power law at $S_{\rm min} = 1$ avoids the scenario of emission dominated by many very faint point sources ($S \ll 1$).

As a check on our understanding of the method, we can increase the area of the signal window and also the number of counts, holding the density of events constant. This change would be expected to have no effect other than reducing the variance in R, since $\sigma(R)$ is expected to scale as $1/\sqrt{\text{number of counts}}$. Fig. 3 demonstrates the improvement in sensitivity.

6.2 Dependence on the PSF and test radius r

The choice of the test radius r significantly impacts the quality of the results. If r is increased sufficiently that the expected number of neighbours for each event is $\gg 1$ then the number of isolated events and empty circles both become small in the diffuse limit and the

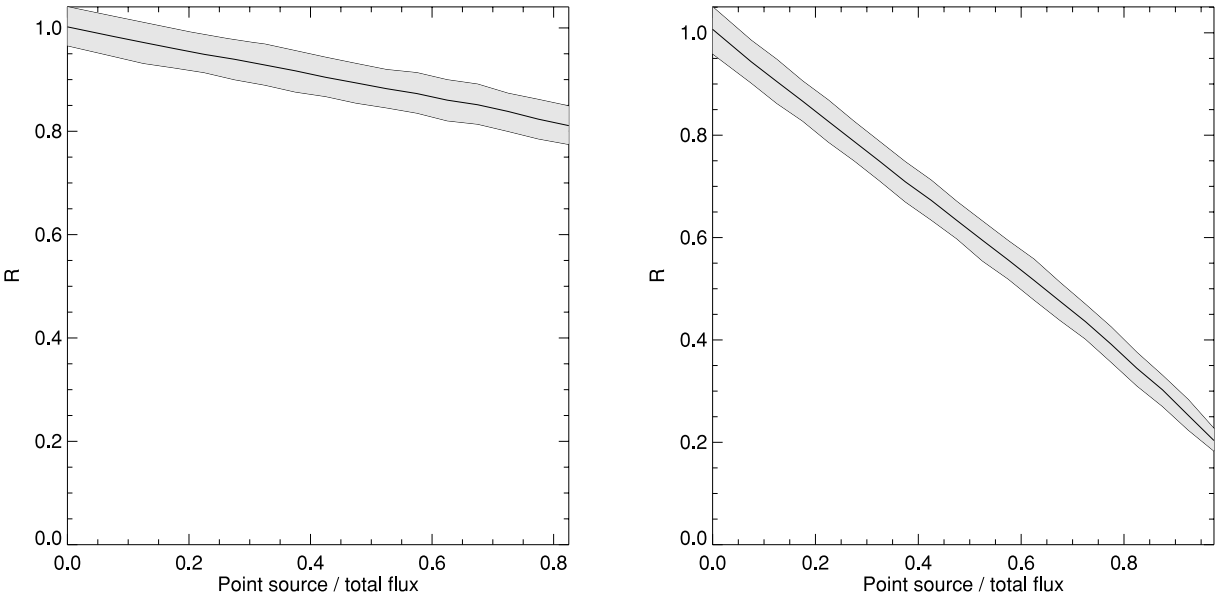

Figure 2. The isotropy ratio R, in Monte Carlo simulated data, in (left) benchmark model 1 and (right-hand panel) benchmark model 2. See Table 1 for definitions of the benchmarks. Lines are 5 and 95 per cent confidence bounds.

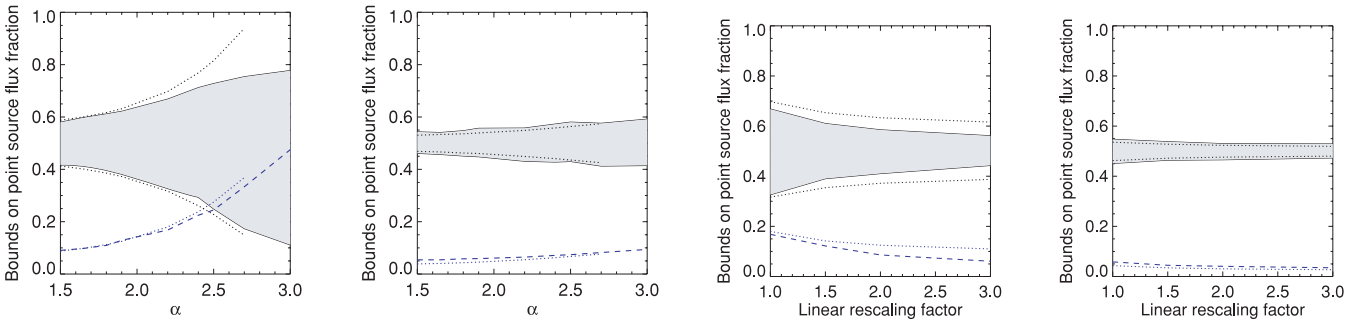

Figure 3. The sensitivity of R as a function of the spectral index of the point-source luminosity function, α (far left: benchmark model 1, centre left: benchmark model 2), and as a function of the signal area while the density of counts is held constant (centre right: benchmark model 1, far right: benchmark model 2). As described in the text, solid lines (bounding the shaded area) indicate the average 90 per cent confidence bounds on the point-source flux fraction from this test when the true fraction is 0.5; the dashed line indicates the 95 per cent confidence upper limit on the point-source flux fraction where $R = 1$. Dotted lines indicate the analogous results for the analytic approximate calculation.

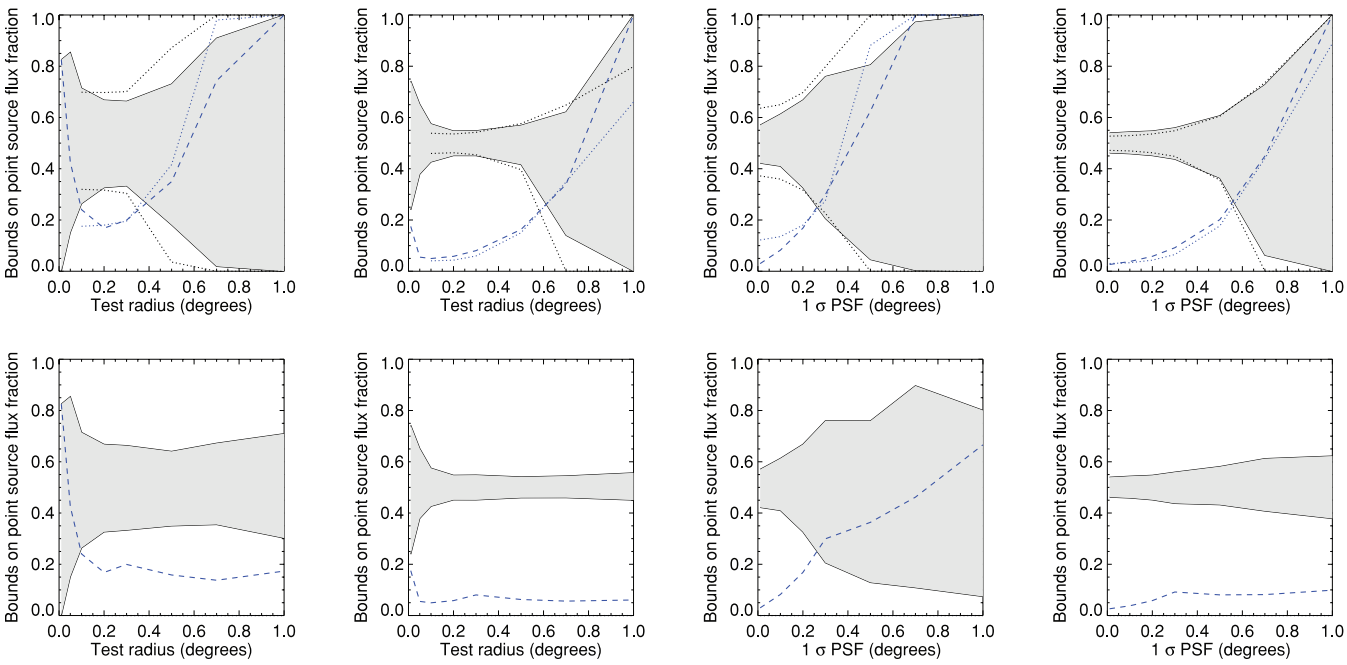

Figure 4. The sensitivity of R as a function of the test radius r, holding the PSF 1σ constant at $0.^{\circ}2$ (far left: benchmark model 1, centre left: benchmark model 2), and as a function of the PSF size, fixing the test radius r to be equal to 1σ for the PSF (centre right: benchmark model 1, far right: benchmark model 2). The top row uses the standard form of the statistic whereas the bottom row uses the modified form (Section 3.1) to prevent the catastrophic failure at large r. As described in the text, solid lines (bounding the shaded area) indicate the average 90 per cent confidence bounds on the point-source flux fraction from this test when the true fraction is 0.5; the dashed line indicates the 95 per cent confidence upper limit on the point-source flux fraction where $R = 1$. In the top row, dotted lines indicate the analogous results for the analytic approximate calculation; note that the truncation approximation made to derive the analytic result breaks down for $r \ll 1\sigma$ PSF.

Poisson fluctuations in R become very large. In the limit where $r \ll$ PSF even events in bright point sources may qualify as 'isolated' and R loses its power to discriminate between point sources and diffuse signal. In the large-r case, a better result with much less noise can be obtained by the generalized method discussed in Section 3.1. These effects are displayed in Fig. 4.

The range of r in which R provides a precise estimate of the point-source flux is given approximately by

$$\mathrm{PSF} \lesssim r \lesssim \sqrt{\frac{\text{area of window}}{\pi \times \text{number of events}}}. \tag{9}$$

If this range is large, r can be varied substantially without much adverse impact on the performance of R as a measure of diffuse flux. For the parameters employed here, the permitted range of r is quite narrow and the effects discussed above are pronounced.

Where this allowed range vanishes, as the average number of counts per PSF circle exceeds 1, this method breaks down as discussed in Section 3, and we should instead employ the generalization discussed in Section 3.1. Fig. 4 shows the slight improvement in the sensitivity of R in the case of a smaller PSF, the breakdown in the method at large PSF and the performance of the generalized method.

6.3 Inhomogeneity in the diffuse flux

The 'diffuse' part of the sample may vary spatially, because of either true spatial variation of the signal or a non-uniform

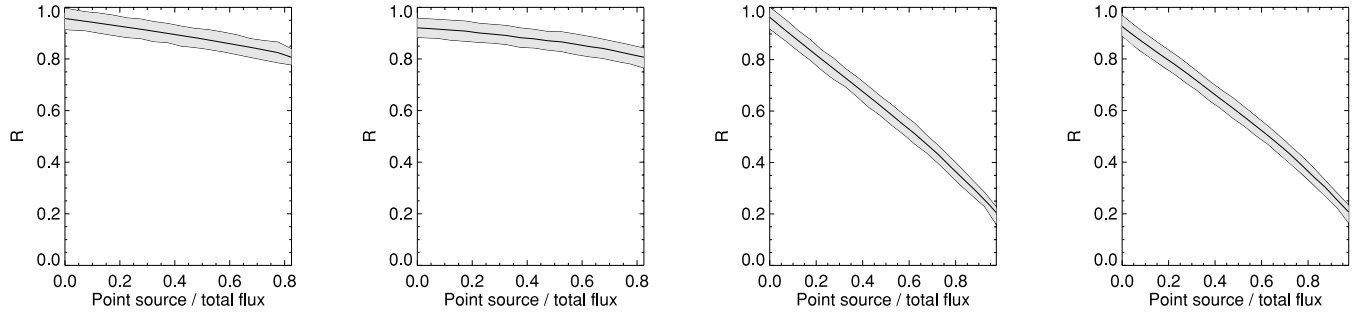

Figure 5. The effect of adding a *b*-dependent tilt to the diffuse photon flux, in Monte Carlo simulated data, for (far left) benchmark model 1, tilt factor 5, (centre left) benchmark model 1, tilt factor 20, (centre right) benchmark model 2, tilt factor 5 and (far right) benchmark model 2, tilt factor 20.

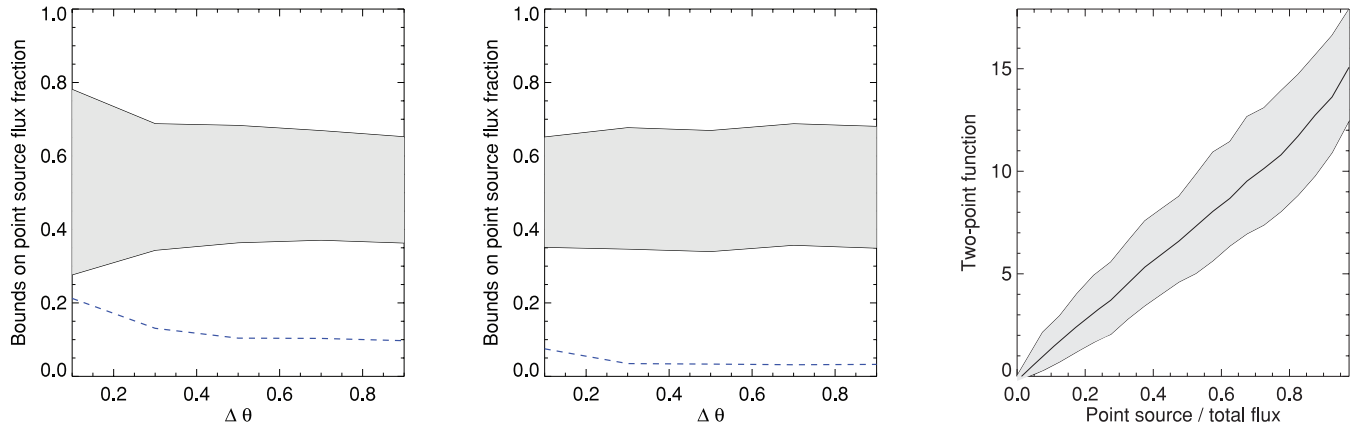

Figure 6. Monte Carlo realizations of the two-point correlation function for the angular region and number of events described in Section 5, varying the annulus width as a fraction of the test radius (the fractional annulus width is denoted $\Delta\theta$) and taking $r = 0.^{\circ}2 = 1\sigma$ of the PSF, for (left-hand panel) benchmark model 1 and (centre panel) benchmark model 2. The right-hand panel shows an example of the mean and 5 and 95 per cent quantiles for the two-point function as a function of the point-source flux fraction, in benchmark model 2 with $\Delta\theta = 0.5$.

instrumental sensitivity or exposure. We show that R is insensitive to such variations by introducing a tilt in the diffuse photon distribution of up to a factor of 20 (i.e. the density of events is 20 times lower at one edge of the signal region than at the other) and observing that it has little effect on R, as shown in Fig. 5. For large tilts, there is a notable downward shift in R in the diffuse limit.

7 COMPARISON TO THE TWO-POINT FUNCTION

In essence, R is a simple measure of the angular correlations between event positions, so it is reasonable to ask how it differs from the two-point correlation function. Consider the case of a single bright point source, compared to two (well-separated) point sources with half the luminosity. The two-point correlation functions for these two situations are quite different, although the total flux from the point sources is the same, because the number of pairs in a given source scales as flux squared. The estimator R, on the other hand, to a first approximation does not probe correlations inside the point sources, and is dependent only on the fraction of diffuse flux, rather than the details of the sources.

This effect tends to reduce the variance of R, relative to the two-point function, as the flux from point sources increases and the luminosity function becomes shallower [i.e. brighter point sources, with (no. of counts)$^2 \gg$ no. of counts, contribute more of the signal]; the R statistic is also less sensitive to the luminosity function parameters than the two-point function. In the limit where the annulus width is large and almost all the non-isolated events have only a single neighbour, i.e. they form a single pair in the calculation of the two-point function (which is the case when the luminosity function is steep and the point-source flux is dominated by faint sources, or when there are simply very few point sources), the two statistics capture essentially the same information and their performance is very similar.

We can directly compare these two statistics as measures of the point-source flux. We employ the unbiased estimator for the two-point function described by Landy & Szalay (1993), and compute its value as a function of the point-source flux contribution (we also scan over the annulus width $\Delta\theta$), within the Monte Carlo framework described previously. Applying the sensitivity measures previously described, the results for the two benchmark parameter sets are shown in Fig. 6. As expected from the discussion above, we see that in the benchmark 1 case, the results are very similar to those previously found, whereas in the more 'optimistic' benchmark 2 case the results are similar in the diffuse limit, but the R test provides much better bounds when the true point-source flux fraction is 0.5.

8 AN EXAMPLE APPLICATION: HIGH ENERGY *Fermi* LAT DATA FROM THE HAZE REGION

The *Fermi* Gamma-ray Space Telescope has recently released all-sky photon data from its first year of operation.[2] The diffuse emission measured by *Fermi* may include signatures of new physics, such as dark matter annihilation or decay, or photons from new

[2]See http://fermi.gsfc.nasa.gov/ssc/data/.

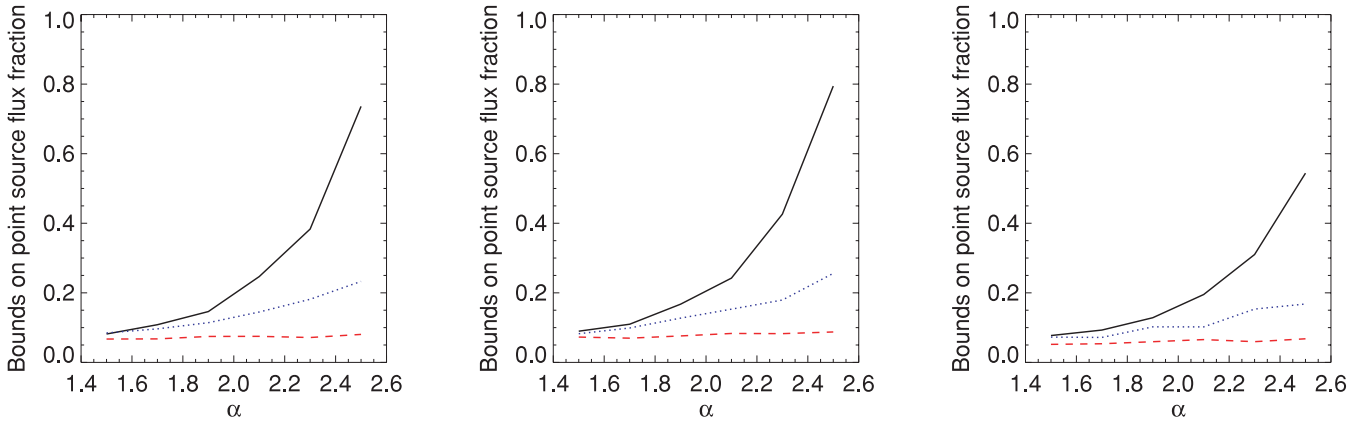

Figure 7. Bounds on the point-source flux fraction in *Fermi* LAT data in the 'Haze' region at 10–100 GeV, from the isotropy ratio R in Monte Carlo simulated data, as a function of the spectral index of the luminosity function. In all cases, $S_{\rm max} = 10$. Left-hand panel: $r = 1\ \sigma$ PSF $= 0.^\circ1$, centre panel: $r = 1\ \sigma$ PSF $= 0.^\circ13$ (the estimated PSF for back-converting events at 10 GeV), right-hand panel: $r = 1\ \sigma$ PSF $= 0.^\circ2$. Black solid line: $S_{\rm min} = 0.01$, blue dotted line: $S_{\rm min} = 0.1$, red dashed line: $S_{\rm min} = 1$.

classes of astrophysical sources. In investigating the origin of the diffuse gamma rays, it will be necessary to estimate what fraction of the observed flux could be due to unresolved point sources. The statistic presented here is well suited to address this question, especially at high energy where the PSF of the *Fermi* LAT is small and the count rate is low.

A microwave excess termed the '*WMAP* Haze' has been observed in the inner 25° of the Galaxy and attributed to synchrotron radiation from some new population of 10–1000 GeV electrons (and/or positrons) (Dobler & Finkbeiner 2008). Recent cosmic ray experiments have also measured a rise in the positron fraction at 10–100 GeV (Adriani et al. 2009), and a hardening in the $e^+ + e^-$ spectrum at 300 GeV – several TeV (Chang et al. 2008; Abdo et al. 2009b), consistent with a new source of high-energy e^+e^-. The *Fermi* LAT can constrain any such new source of electrons by searching for gamma rays from inverse Compton scattering of the electrons on starlight.

As an example of how this statistic can be applied, we consider the Class 3 (diffuse class) events measured by the *Fermi* LAT with energies between 10 and 100 GeV, in the region of the sky optimized for study of the *WMAP* Haze, defined in Galactic coordinates by $|l| < 15$, $-30 < b < -10$. There are 1146 such events in this signal window. We take the signal region $|l| < 18$, $-40 < b < -8$. Above 10 GeV, the 68 per cent containment radius of the LAT is $\lesssim 0.^\circ2$ (Rando & the Fermi LAT Collaboration 2009), which corresponds to a PSF of $\sim 0.^\circ13$ in the sense that we have used (1D σ for the Gaussian distribution of photons from a single point source). We consider three values of the PSF: the best-estimate upper bound of $0.^\circ13$, and also $0.^\circ1$ and $0.^\circ2$, to demonstrate the possible effect of uncertainties in the PSF, or a varying PSF over the energy range of interest. In all cases, we take the test radius r to be equal to the PSF and generate 10^6 random points to determine n_E.

Fig. 7 shows the 95 per cent confidence limits on the fraction of flux originating from point sources, as a function of $S_{\rm min}$ and the luminosity function spectral index α. We see that even with fairly pessimistic assumptions for the luminosity function, as in the 'Benchmark 1' case, the measured values of R in the Fermi data are outside the 95 per cent confidence limits if the flux from sources with >0.1 (expected) counts $\rm yr^{-1}$ exceeds ∼15 per cent of the total; the limits can be significantly stronger if the assumed point-source luminosity function is shallower or extends to higher flux. Note also that we have *not* subtracted resolved point sources in this region; there are no known point sources in this region of the sky in the *Fermi* 3-month bright source list.[3]

9 CONCLUSION

We have introduced a simple and easily calculable statistic that linearly traces the fraction of flux arising from diffuse emission, as opposed to unresolved point sources. The statistic is quite insensitive to even pronounced large-scale anisotropies in the diffuse emission, such as might originate from the proximity of a bright region or angular variation in the detector exposure. The linear response of this statistic to flux originating from point sources, and its smaller variance, makes it superior to the two-point correlation function as a tracer of emission from unresolved point sources.

The sensitivity of the statistic to point-source emission naturally depends on the luminosity function of the point sources, as a sufficiently steep power law extending to sufficiently small luminosities is strictly indistinguishable from diffuse emission. However, the statistic retains discriminatory power for spectral indices up to $\alpha \sim 3$, with a low-luminosity cut-off corresponding to an average of 0.1 counts, and assuming all point sources with average luminosity $\gtrsim$10 counts are resolved and removed. Known luminosity functions for astrophysical point sources generically have shallower slopes than this limit.

When the average number of events per PSF circle exceeds 1, the original form of the statistic breaks down: however, we have described a simple generalization suitable for this case, and demonstrated its efficacy. Increasing the number of counts by taking additional sky regions into account (i.e. without increasing the density of events) improves the variance by the usual $1/N_{\rm event}$ Poisson factor. This statistic generalizes readily to higher dimensions; possible applications include the study of void statistics (Fry 1986).

As an example, we have applied this statistic to Class 3 (diffuse class) photon data from the *Fermi* LAT in the angular region relevant for study of the *WMAP* Haze, at energies of 10–100 GeV. We find that even with rather pessimistic assumptions for the point-source luminosity function, at most ∼15 per cent of the emission in this region can be attributed to unresolved point sources with average

[3] http://fermi.gsfc.nasa.gov/ssc/data/access/lat/bright_src_list/

luminosities of 0.1+ counts yr^{-1}, and the results are consistent with 100 per cent diffuse emission.

ACKNOWLEDGMENTS

We wish to acknowledge helpful conversations with Marc Davis, Josh Grindlay, Igor Moskalenko, Jim Peebles and Pat Slane. We thank the anonymous referee for helpful comments. TRS is supported by a Sir Keith Murdoch Fellowship from the American Australian Association.

REFERENCES

Abdo A. A. et al., 2009a, ApJ, 700, 597
Abdo A. A. et al., 2009b, Phys. Rev. Lett., 102, 181101
Abraham J. et al., 2007, Sci, 318, 938
Abraham J. et al., 2008, Astropart. Phys., 29, 188
Adriani O. et al., 2009, Nat, 458, 607
Ave M. et al., 2009, J. Cosmology Astropart. Phys., 07, 023
Chang J. et al., 2008, Nat, 456, 362
Dobler G., Finkbeiner D. P., 2008, ApJ, 680, 1222
Fry J. N., 1986, ApJ, 306, 358
Gilfanov M., 2004, MNRAS, 349, 146
Grimm H. J., Gilfanov M., Sunyaev R., 2003, MNRAS, 339, 793
Kashlinsky A., Mather J. C., Odenwald S., 1996, ApJ, 473, 9
Kim D.-W., Fabbiano G., 2004, ApJ, 611, 846
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Padovani P., Giommi P., Landt H., Perlman E. S., 2007, ApJ, 662, 182
Rando R., the Fermi LAT Collaboration 2009, preprint (arXiv:0907.0626)
Scheuer P. A. G., 1974, MNRAS, 166, 329
Voss R., Gilfanov M., 2006, A&A, 447, 71
Voss R. et al., 2009, ApJ, 701, 471

APPENDIX A: THE ANALYTIC GRID MODEL AND THE LUMINOSITY FUNCTION

The analytic estimates for R and $\sigma(R)$ derived in Section 3 are functions of the fraction of cells which gain exactly 0–1 counts from point sources. If we are to compare the analytic estimates to the results of the MC runs, the S and T parameters must be expressed in terms of α, $S_{\min}$ and $S_{\max}$. This is a nontrivial exercise, and in any case the analytic estimates and the MC runs should not be expected to agree in detail, since they use different criteria for determining neighbouring events (events within the same cell versus events within the test radius). The main purpose of the analytic estimates is to demonstrate the general scaling behaviour of R and $\sigma(R)$, and the regions of parameter space where this test loses discriminatory power.

None the less, for completeness, we now derive approximate relations between S and T and the parameters of the luminosity function. These relations are used to provide an analytic estimate for the sensitivity of R as a measure of the point-source flux fraction, in Figs 3–4. Here, we only derive the *mean* values for S and T for a given luminosity function; for any choice of the luminosity function, there will be an additional contribution to $\sigma^2(R)$ from the variances of S and T, which is not taken into account in this analysis.

Let us first compute the probabilities for a given (i,j) cell to obtain exactly one or zero counts from the addition of a single (randomly placed) point-source, denoted ρ_1^{ij} and ρ_0^{ij}, respectively. If the point source has an expected contribution of k counts, the PSF is assumed to be Gaussian with variance σ^2, the cells are square with dimensions $\delta \times \delta$, and the point source is centred at (x_0,y_0), then the expected number of counts in a cell with left-hand lower corner (x_i, y_j) is given by,

$$\begin{aligned}\lambda_{ij}(k, x_0, y_0) &= \frac{k}{2\pi\sigma^2}\int_{x_i}^{x_i+\delta}\int_{y_j}^{y_j+\delta} \mathrm{e}^{-(x-x_0)^2/2\sigma^2}\mathrm{e}^{-(y-y_0)^2/2\sigma^2}\mathrm{d}x\mathrm{d}y,\\ &= \frac{k}{4}\left[\mathrm{Erf}\left(\frac{x_0-x_i}{\sqrt{2}\sigma}\right)-\mathrm{Erf}\left(\frac{x_0-x_i-\delta}{\sqrt{2}\sigma}\right)\right]\\ &\quad\times\left[\mathrm{Erf}\left(\frac{y_0-y_j}{\sqrt{2}\sigma}\right)-\mathrm{Erf}\left(\frac{y_0-y_j-\delta}{\sqrt{2}\sigma}\right)\right].\end{aligned} \tag{A1}$$

The Poisson probabilities to obtain exactly one and zero counts in this cell, from a point source providing k events, are then given by

$$\begin{aligned}\rho_1^{ij}(k, x_0, y_0) &= \lambda_{ij}(k, x_0, y_0)\mathrm{e}^{-\lambda_{ij}(k,x_0,y_0)},\\ \rho_0^{ij}(k, x_0, y_0) &= \mathrm{e}^{-\lambda_{ij}(k,x_0,y_0)}.\end{aligned} \tag{A2}$$

Since the λ_{ij} depend only linearly on k, it is straightforward to integrate over the luminosity function: writing $\lambda_{ij} = k\theta$, we obtain

$$\begin{aligned}\rho_0^{ij}(x_0, y_0) &= \frac{\int_{S_{\min}}^{S_{\max}} k^{-\alpha}\mathrm{e}^{-k\theta}\mathrm{d}k}{\int_{S_{\min}}^{S_{\max}} k^{-\alpha}\mathrm{d}k}\\ &= \frac{\theta^{\alpha-1}\left(\Gamma\left(1-\alpha, S_{\min}\theta\right)-\Gamma\left(1-\alpha, S_{\max}\theta\right)\right)(1-\alpha)}{S_{\max}^{1-\alpha}-S_{\min}^{1-\alpha}},\end{aligned} \tag{A3}$$

$$\begin{aligned}\rho_1^{ij}(x_0, y_0) &= \frac{\int_{S_{\min}}^{S_{\max}} k^{-\alpha}k\theta\mathrm{e}^{-k\theta}\mathrm{d}k}{\int_{S_{\min}}^{S_{\max}} k^{-\alpha}\mathrm{d}k}\\ &= \frac{\theta^{\alpha-1}\left[\Gamma\left(2-\alpha, S_{\min}\theta\right)-\Gamma\left(2-\alpha, S_{\max}\theta\right)\right](1-\alpha)}{S_{\max}^{1-\alpha}-S_{\min}^{1-\alpha}}.\end{aligned} \tag{A4}$$

Strictly speaking, we should now integrate this result with respect to x_0 and y_0; however, this is not analytically tractable. Provided σ is not too much smaller than δ, it is a good approximation to instead integrate equation (A1) over x_0 and y_0 within a given cell, obtaining an average expected number of counts $\lambda_{ij}(k) = k\theta$ for each cell, and then use this result for θ in equation (A4). Let us choose our coordinate system so that (x_0, y_0) lies in a cell with left-hand corner $(0, 0)$ and the ij cell has left-hand corner $(i\delta, j\delta)$. For $|i|$ or $|j| \gg 0$, λ_{ij} will be negligible, so we can make the further approximation of truncating the sum over i, j at some point. In this work, we make the approximation that for cells with $|i|$ or $|j| > 2$, $\rho_0^{ij} = 1$ and $\rho_1^{ij} = 0$.

We can then write θ for each cell in terms of three functions of σ/δ, denoted t, u and v, obtained by averaging $(1/2)[\mathrm{Erf}(\frac{x_0-x_i}{\sqrt{2}\sigma}) - \mathrm{Erf}(\frac{x_0-x_i-\delta}{\sqrt{2}\sigma})]$ over $x_0 = [0, \delta]$, for $i = 0, 1, 2$, respectively:

$$\begin{aligned}t &= \mathrm{Erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)-\sqrt{\frac{2}{\pi}}\frac{\sigma}{\delta}\left(1-\mathrm{e}^{-\delta^2/2\sigma^2}\right),\\ u &= \frac{\sigma}{\delta\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{2\delta^2}{\sigma^2}}-2\mathrm{e}^{\frac{\delta^2}{2\sigma^2}}+1\right)-\mathrm{Erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)+\mathrm{Erf}\left(\frac{\sqrt{2}\delta}{\sigma}\right),\\ v &= \frac{1}{2}\left[\mathrm{Erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)+3\mathrm{Erf}\left(\frac{3\delta}{\sqrt{2}\sigma}\right)-4\mathrm{Erf}\left(\frac{\sqrt{2}\delta}{\sigma}\right)\right]\\ &\quad+\frac{\sigma}{\delta\sqrt{2\pi}}\left(\mathrm{e}^{-\frac{9\delta^2}{2\sigma^2}}-2\mathrm{e}^{-2\delta^2/\sigma^2}+\mathrm{e}^{-\delta^2/2\sigma^2}\right).\end{aligned} \tag{A5}$$

In terms of these functions, the values of θ for the relevant cells are

$i = j = 0$, one cell, $\theta = t^2$

$i = 0,\ j = \pm 1$ or $i = \pm 1,\ j = 0$, four cells, $\theta = tu$

$i = \pm 1,\ j = \pm 1, \quad$ four cells, $\quad \theta = u^2$

$i = 0,\ j = \pm 2 \quad$ or $\quad i = \pm 2,\ j = 0, \quad$ four cells, $\quad \theta = tv$

$i = \pm 1,\ j = \pm 2 \quad$ or $\quad i = \pm 2,\ j = \pm 1, \quad$ eight cells, $\quad \theta = uv$

$i = \pm 2,\ j = \pm 2, \quad$ four cells, $\quad \theta = v^2$.

Summing over the probabilities ρ_0^{ij} (ρ_1^{ij}) for all cells then yields the expected number of cells which gain zero (one) counts from the addition of a point source, denoted E_0 (E_1). If n sources are added, the probability of any one cell gaining zero counts is $(E_0/N)^n$, and the probability of a single count is $n(E_0/N)^{n-1}$ (E_1/N) (i.e. one source contributes a single count, all others contribute zero). Multiplying by the number of cells N yields the $N - T$ and S parameters, respectively. The number of sources n is related to the average total emission from point sources,

$$n = \frac{2-\alpha}{1-\alpha}\left(\frac{S_{\rm max}^{1-\alpha} - S_{\rm min}^{1-\alpha}}{S_{\rm max}^{2-\alpha} - S_{\rm min}^{2-\alpha}}\right) \times \text{mean total counts}. \tag{A6}$$

In order to compare the analytic result (based on the grid) to the MC results (based on neighbours within a test radius), we must also impose a relation between the side length of the grid cells and the test radius. In Figs 3–4, we require that the area of a grid cell is the same as the area of a PSF circle, i.e. $\delta = \sqrt{\pi} r$; a different prescription might give better agreement between the MC results and the estimates from the grid model.

This paper has been typeset from a TEX/LATEX file prepared by the author.